PTO/SB/21 (11-07)
Approved for use through 11/30/2007. OMB 0651-0031
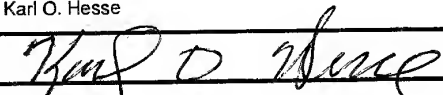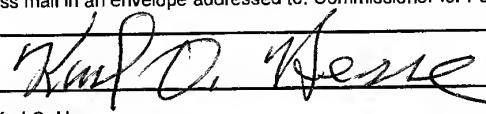U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

TRANSMITTAL FORM

(to be used for all correspondence after initial filing)

Total Number of Pages in This Submission | 48

Application Number	10/629,448
Filing Date	July 29, 2003
First Named Inventor	Kelkar et. al.
Art Unit	1631
Examiner Name	Loria Clow
Attorney Docket Number	CHA9 2003 0003 US1

ENCLOSURES   *(Check all that apply)*

- [ ] Fee Transmittal Form
  - [ ] Fee Attached
- [ ] Amendment/Reply
  - [ ] After Final
  - [ ] Affidavits/declaration(s)
- [ ] Extension of Time Request
- [ ] Express Abandonment Request
- [ ] Information Disclosure Statement
- [ ] Certified Copy of Priority Document(s)
- [ ] Reply to Missing Parts/ Incomplete Application
  - [ ] Reply to Missing Parts under 37 CFR 1.52 or 1.53

- [ ] Drawing(s)
- [ ] Licensing-related Papers
- [ ] Petition
- [ ] Petition to Convert to a Provisional Application
- [ ] Power of Attorney, Revocation Change of Correspondence Address
- [ ] Terminal Disclaimer
- [ ] Request for Refund
- [ ] CD, Number of CD(s) _____
  - [ ] Landscape Table on CD

- [ ] After Allowance Communication to TC
- [ ] Appeal Communication to Board of Appeals and Interferences
- [ ] Appeal Communication to TC **(Appeal Notice, Brief, Reply Brief)**
- [ ] Proprietary Information
- [ ] Status Letter
- [x] Other Enclosure(s) (please Identify below):

Resubmission of Appeal Brief previously filed on October 3, 2007.

Remarks

It is believed that no additional fees are due at this time, however, in the event that an additional fee is required, please charge that fee to deposit account number 09-0469.

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT

Firm Name	Karl O. Hesse
Signature	*[signature]*
Printed name	Karl O. Hesse
Date	November 7, 2007
Reg. No.	25,398

CERTIFICATE OF TRANSMISSION/MAILING

EFS WEB

I hereby certify that this correspondence is being ~~facsimile transmitted~~ to the USPTO or deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on the date shown below:

Signature	*[signature]*
Typed or printed name	Karl O. Hesse
Date	November 7, 2007

Date: November 7, 2007

In re application of:

Kelkar et al

Serial No.: 10/629,448

Filed: July 29, 2003

Group Art Unit: 1631

Examiner: **Loria Clow**

FOR: **Method and Program**

Product for Discovering

Similar Gene Expression Profiles

AMENDED APPEAL BRIEF IN SUPPORT OF APPEAL FROM
THE PRIMARY EXAMINER TO THE BOARD OF APPEALS

Assistant Commissioner for Patents

Washington DC 20231

Sir:

Appellants herewith submit an Amended Appeal Brief in support of the appeal to the Board of Patent Appeals and Interferences from the decision dated May 16, 2007 of the Primary Examiner finally rejecting claims 1-6, 10-16 and 20.

(I) Real Party in Interest

The real party in interest in this appeal is International Business Machines Corporation, a New York corporation, assignee of the entire right, title and interest in the claimed invention.

(II) Related Appeals and Interferences

No other appeals or interferences are known to the Appellants, the Appellants' legal representative, or assignee that will directly affect or be directly affected by or have a bearing on the Board's decision in this appeal.

(III) Status of Claims

Claims 1-6, 10-16 and 20 are pending in this application.
Claims 7-9 and 17-19 were canceled after restriction.
The rejection of claims 10-16 and 20 under 35 U.S.C. 101 and
the rejection of claims 1-6 under 35 U.S.C. 101 and
for new matter is appealed.

When this application was filed in 2003, applicants' attorney believed that providing independent claims would facilitate prosecution because they could be allowed or rejected without requiring rewriting to incorporate independent claim limitations into dependent claims when they were found to be allowable. It is understood that currently this practice is not permitted in excess of four claims and applicants' attorney regrets any inconvenience caused thereby.

(IV) Status of Amendments

The amendment filed before final has been entered.
The amendment filed after final has not been entered.

(V) Summary of Claimed Subject Matter

References to paragraphs are made to the application as published.

1. A method for determining similarity between portions of gene expression profiles in a computer comprising the steps of:

processing a number of gene expression profiles with a similar sequences algorithm that is a time and intensity invariant correlation function to obtain a data set of gene expression profile pairs and a match fraction for each gene expression profile pair; **Page 7,Line 31-Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**

listing gene expression profile pairs in clusters by their match fractions; **Page 8 Lines 16-26 (paragraph 38) and Fig.2 # 223**

removing a first gene expression profile from a cluster when another cluster has another gene expression profile with a higher match fraction with the first gene expression profile, unless the another gene expression profile requires a larger number of subsequences to achieve similarity with the first gene expression profile; **Page 8 Line 27-Page 9 Line 7 (paragraph 39) and Fig. 2 # 225**

repeating the removing step until all gene expression profiles are listed in only one cluster; **Page 8 Line 27-Page 9 Line 7 (Paragraph 39) last sentence**

providing output of the listing of clusters of gene expression profiles. **Page 6 Lines 24-31 and Page 7 Lines 19-24 (Paragraphs 32 and 34) and Fig. 8 #s 54 and 66**

2. A method for determining similarity between portions of gene expression profiles comprising the steps of:

processing a number of gene expression profiles with a similar sequences algorithm that is a time and intensity invariant correlation function to obtain a data set of gene expression pairs and a match fraction for each pair; **Page 7,Line 31 - Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215)**

listing gene expression pairs in clusters by their match fractions; **Page 8 Lines 16-26 (Paragraph 38) and Fig.2 # 223**

removing a first gene from a first cluster when the first gene is also in a second cluster which has another gene with a higher match fraction with the first gene than any of the genes in the first cluster have with the first gene, but; **Page 8 Line 27 - Page 9 Line 7 (Paragraph 39) and Fig. 2 # 225**

retaining the first gene in the first cluster and removing the first gene from the second cluster when the difference between the highest match fraction of the first gene with a gene in the first cluster and the highest match fraction of the first gene with a gene in the second cluster is less than a minimum difference threshold and the number of subsequences represented in the similar gene pair having the highest match fraction in the first cluster is higher than the number of subsequences represented in the similar gene pair having the highest match fraction in the second cluster; **Page 9 Line 18 - Page 10 Line 10 (Paragraphs 41, 42, 43 and 56) and Fig. 2 # 225**

repeating the removing step until all genes are listed in only one cluster; **Page 8 Line 27 - Page 9 Line 7 (Paragraph 39) last sentence**

providing output of the listing of clusters of gene expression profiles. **Page 6 Lines 24 - 31 and Page 7 Lines 19-24 (Paragraphs 32 and 34) and Fig. 8 #s 54 and 66**

1   3. A method of determining similarity between portions of gene
2 expression profiles comprising the steps of:
3   processing data embodying a number of gene expression profiles
4 with a similar sequences algorithm in a computer that is a time and
5 intensity invariant correlation function to obtain a data set of
6 gene expression pairs and a match fraction for each pair; **Page
7 7,Line 31 - Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**
8   choosing a threshold match fraction; **Page 13 Lines 6-20**
9 **(Paragraph 52) sentence 6 and Fig.2 # 217**
10   listing gene expression pairs in clusters by their match
11 fractions above the threshold; **Page 13 Lines 15-20 (Paragraph 52)**
12 **and Fig.2 #s 221, 215, 219 and 223**
13   adding each gene not already in a cluster to a cluster having
14 another gene having a highest match fraction with the each gene
15 without regard of the threshold; **Page 9 Lines 8-14 (Paragraph 40)**
16 **and Fig 2 # 229**
17   removing a first gene from a cluster when the first gene is
18 also in another cluster which has another gene with a higher match
19 fraction with the first gene than any of the genes in the cluster
20 have with the first gene; **Page 8 Line 27-Page 9 Line 7 (Paragraph**
21 **39) first sentence and Fig. 2 # 225**
22   repeating the removing step until all genes are listed in only
23 one cluster; **Page 8 Line 27-Page 9 Line 7 (Paragraph 39) last**
24 **sentence**
25   providing output of the listing of clusters of gene
26 expression profiles. **Page 6 Lines 24-31 and Page 7 Lines 19-24**
27 **(Paragraphs 32 and 34) and Fig. 8 #s 54 and 66**

1   4. A method for determining similarity between portions of gene
2 expression profiles comprising the steps of:
3   processing a number of gene expression profiles with a similar
4 sequences algorithm that is a time and intensity invariant
5 correlation function with a computer to obtain a data set of gene
6 expression pairs and a match fraction for each pair; **Page 7,Line 31-**
7 **Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**
8   choosing a threshold match fraction; **Page 13 Lines 5-20**
9 **(Paragraph 52) sentence 6 and Fig.2 # 217**
10   listing gene expression pairs in clusters by their match
11 fractions above the threshold; **Page 8 Lines 16-26 (paragraph 38) and**
12 **Fig.2 # 223**
13   adding each gene not already in a cluster to a cluster having
14 another gene having a highest match fraction disregarding the
15 threshold with the each gene; **Page 9 Lines 8-14 (Paragraph 40) and**
16 **Fig 2 # 229**
17   removing a first gene from a first cluster when the first gene
18 is also in a second cluster which has another gene with a higher
19 match fraction with the first gene than any of the genes in the
20 first cluster have with the first gene, **Page 8 Line 27-Page 9 Line 7**
21 **(Paragraph 39) first sentence and Fig. 2 # 225**
22 but;
23   retaining the first gene in the first cluster and removing the
24 first gene from the second cluster when the difference between the
25 highest match fraction of the first gene with a gene in the first
26 cluster and the highest match fraction of the first gene with a gene
27 in the second cluster is less than a minimum difference threshold
28 and the number of subsequences represented in the similar gene pair
29 having the highest match fraction in the first cluster is higher
30 than the number of subsequences represented in the similar gene pair

having the highest match fraction in the second cluster; **Page 8 Line 27-Page 9 Line 7 (Paragraph 39) second sentence and Fig. 2 # 225**

repeating the removing and retaining steps until all genes are listed in only one cluster; **Page 8 Line 27-Page 9 Line 7 (Paragraph 39) last sentence**

providing output of the listing of clusters of gene expression profiles. **Page 6 Lines 24-31 and Page 7 Lines 19-24 (Paragraphs 32 and 34) and Fig. 8 #s 54 and 66**

5. A method in a computer for determining similarity between genes comprising the steps of:

listing genes to be compared in a data set by their gene expression profiles; **Page 7,Line 31-Page 8,Line 9 (Paragraph 36) second sentence and Fig. 1 # 211**

processing the listed gene expression profiles with a similar sequences algorithm that is a time and intensity invariant correlation function to obtain a data set of gene expression pairs and a match fraction for each pair; **Page 7,Line 31-Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**

choosing a threshold match fraction; **Page 13 Lines 5-20 (Paragraph 52) sentence 6 and Fig.2 # 217**

creating a set G in which to list indices of genes accounted for; **Page 8 Lines 11-14 (Paragraph 37) and Fig. 2 # 217**

assigning genes i and j to a cluster a if they have a match fraction greater than the threshold; **Page 8 Lines 16-26 (Paragraph 38) second sentence and Fig. 2 # 223 -> ca={i,j,..}**

assigning gene k to the cluster a if it has a match fraction greater than the threshold with either gene i or gene j; **Page 8 Lines 16-26 (Paragraph 38) last sentence and Fig. 2 # 223 -> ca={i,j,k,..}**

assigning genes k and l to a cluster b if they have a match fraction greater than the threshold and if both gene k and gene l do not have match fractions above the threshold with either gene i or gene j; **Page 8 Lines 16-26 (Paragraph 38) and Fig. 2 # 223 -> cb={k,l,..}**

repeating the assigning steps until all genes to be compared have been considered; **Page 13 Lines 5-20 (Paragraph 52) first two sentences and Fig #s 217, 219, and 229**

removing a first gene from a cluster when another cluster has another gene with a higher match fraction with the first gene; **Page 8 Line 27-Page 9 Line 7 (Paragraph 39) second sentence and Fig. 2 # 225**

repeating the removing step until all genes are listed in only one cluster; **Page 8 Line 27-Page 9 Line 7 (Paragraph 39) last sentence**

providing output of the listing of clusters of gene expression profiles. **Page 6 Lines 24-31 and Page 7 Lines 19-24 (Paragraphs 32 and 34) and Fig. 8 #s 54 and 66**

1   6. A method in a computer for determining similarity between
2 genes comprising the steps of:
3   listing genes ·to be compared in a data set by their gene
4 expression profiles; **Page 7,Line 31-Page 8,Line 9 (Paragraph 36)**
5 **second sentence and Fig. 1 # 211**
6   processing the listed gene expression profiles with a similar
7 sequences algorithm that is a time and intensity invariant
8 correlation function to obtain a data set of gene expression pairs
9 and a match fraction for each pair; **Page 7,Line 31-Page 8,Line 9**
10 **(Paragraph 36) and Fig. 1, # 213, 215**
11   choosing a threshold match fraction; **Page 13 Lines 5-20**
12 **(Paragraph 52) sentence 6 and Fig.2 # 217**
13   creating a set G in which to list indices of genes accounted
14 for; **Page 13 Lines 5-20 (Paragraph 52) 3rd and 4th sentences and**
15 **Fig. 2 # 217**
16   assigning genes i and j to cluster 1 if they have a match
17 fraction greater than the threshold; **Page 13 Line 21-Page 14 Line 6**
18 **and Page 14 Lines 8-12 (Paragraphs 53 and 54) first sentence and**
19 **Table IV**
20   assigning gene k to cluster 1 if it has a match fraction
21 greater than the threshold with either gene i or gene j; **Page 13**
22 **Line 21-Page 14 Line 6, Page 14 Lines 8-12, Page 14 Line 24-Page 15**
23 **Line 3 (Paragraphs 53, 54 2nd sentence, 56) and Table IV**
24   assigning genes k and l to cluster 2 if they have a match
25 fraction greater than the threshold and if both gene k and gene l do
26 not have match fractions above the threshold with either gene i or
27 gene j; **Page 13 Line 21-Page 14 Line 6, Page 14 Lines 8-12, Page 14**
28 **Line 24-Page 15 Line 3 (Paragraphs 53, 54 2nd sentence, 56) and**
29 **Table IV and Fig. 2 # 225**
30   removing a first gene from a cluster when another cluster has
31 another gene with a higher match fraction with the first gene,

unless the another gene requires a larger number of subsequences to achieve similarity with the first gene; **Page 14 Line 24 – Page 15 Line 3 (Paragraph 56) beginning at the 3rd sentence and Fig. 2 # 225**

    repeating the removing step until all genes are listed in only one cluster; **Page 8 Line 27-Page 9 Line 7 (Paragraph 39) last sentence**

    providing output of the listing of clusters of gene expression profiles. **Page 6 Lines 24-31 and Page 7 Lines 19-24 (Paragraphs 32 and 34) and Fig. 8 #s 54 and 66**

1   10. A program product having computer readable code stored on a
2 recordable media for determining similarity between portions of gene
3 expression profiles comprising:
4   programmed means for processing a number of gene expression
5 profiles with a similar sequences algorithm that is a time and
6 intensity invariant correlation function to obtain a data set of
7 gene expression pairs and a match fraction for each pair;
8 **Page 7,Line 31-Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**
9   programmed means for listing gene expression pairs in clusters
10 by their match fractions; **Page 8 Lines 16-26 (paragraph 38) and**
11 **Fig.2 # 223**
12   programmed means for removing a first gene from a cluster when
13 the first gene is also in another cluster which has another gene
14 with a higher match fraction with the first gene than any of the
15 genes in the cluster have with the first gene; **Page 8 Line 27-Page 9**
16 **Line 7 (paragraph 39) and Fig. 2 # 225**
17   programmed means for repeating the removing step until all
18 genes are listed in only one cluster. **Page 8 Line 27-Page 9 Line 7**
19 **(Paragraph 39) last sentence**

1   11. A program product having computer readable code stored on a
2   recordable media for determining similarity between portions of gene
3   expression profiles using output from a similar sequences algorithm
4   that is a time and intensity invariant correlation function
5   comprising:
6   programmed means for providing a gene expression profile data
7   set as input to programmed means embodying a similar sequences
8   algorithm that is a time and intensity invariant correlation
9   function to obtain a data set of gene expression pairs and a match
10   fraction for each pair as output from the programmed means embodying
11   a similar sequences algorithm; **Page 7,Line 31-Page 8,Line 9; Page**
12   **10, Line 22-Page 11,Line 7 (Paragraphs 36, 44) and Fig. 1, # 213,**
13   **215)**
14   programmed means for listing the gene expression pairs in
15   clusters by their match fractions; **Page 7,Line 31-Page 8,Line 9**
16   **(Paragraph 36) beginning at the second sentence and Fig. 1 # 215**
17
18   programmed means for removing a first gene from a cluster when
19   the first gene is also in another cluster which has another gene
20   with a higher match fraction with the first gene than any of the
21   genes in the cluster have with the first gene; **Page 8 Line 27-Page 9**
22   **Line 7 (Paragraph 39) first sentence and Fig. 2 # 225**
23   programmed means for repeating the removing step until all
24   genes are listed in only one cluster. **Page 8 Line 27-Page 9 Line 7**
25   **(Paragraph 39) last sentence**

1   12. A program product having computer readable code stored on a
2   recordable media for determining similarity between portions of gene
3   expression profiles comprising the steps of:
4   programmed means for processing a number of gene expression
5   profiles with a similar sequences algorithm that is a time and
6   intensity invariant correlation function to obtain a data set of
7   gene expression pairs and a match fraction for each pair; **Page**
8   **7,Line 31-Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**
9   programmed means for listing gene expression pairs in clusters
10 by their match fractions; **Page 8 Lines 16-26 (paragraph 38) and**
11 **Fig.2 # 223**
12   programmed means for removing a first gene from a first cluster
13 when the first gene is also in a second cluster which has another
14 gene with a higher match fraction with the first gene than any of
15 the genes in the first cluster have with the first gene, **Page 8 Line**
16 **27-Page 9 Line 7 (Paragraph 39) first sentence and Fig. 2 # 225**
17 but;
18   programmed means for retaining the first gene in the first
19 cluster and removing the first gene from the second cluster when the
20 difference between the highest match fraction of the first gene with
21 a gene in the first cluster and the highest match fraction of the
22 first gene with a gene in the second cluster is less than a minimum
23 difference threshold and the number of subsequences represented in
24 the similar gene pair having the highest match fraction in the first
25 cluster is higher than the number of subsequences represented in the
26 similar gene pair having the highest match fraction in the second
27 cluster; **Page 8 Line 27-Page 9 Line 7 (Paragraph 39) second sentence**
28 **and Fig. 2 # 225**
29   programmed means for repeating the removing step until all
30 genes are listed in only one cluster. **Page 8 Line 27-Page 9 Line 7**
31 **(Paragraph 39) last sentence**

1   13. A program product having computer readable code stored on a
2 recordable media for determining similarity between portions of gene
3 expression profiles comprising the steps of:
4   programmed means for processing a number of gene expression
5 profiles with a similar sequences algorithm that is a time and
6 intensity invariant correlation function to obtain a data set of
7 gene expression pairs and a match fraction for each pair; **Page**
8 **7,Line 31-Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**
9   programmed means for choosing a threshold match fraction; **Page**
10 **13 Lines 6-20 (Paragraph 52) sentence 6 and Fig.2 # 217**
11   programmed means for listing gene expression pairs in clusters
12 by their match fractions above the threshold; **Page 8 Lines 16-26**
13 **(paragraph 38) and Fig.2 # 223**
14   programmed means for adding each gene not already in a cluster
15 to a cluster having another gene having a highest match fraction
16 with the each gene without regard of the threshold; **Page 9 Lines 8-**
17 **14 (Paragraph 40) and Fig 2 # 229**
18   programmed means for removing a first gene from a cluster when
19 the first gene is also in another cluster which has another gene
20 with a higher match fraction with the first gene than any of the
21 genes in the cluster have with the first gene; **Page 8 Line 27-Page 9**
22 **Line 7 (Paragraph 39) first sentence and Fig. 2 # 225**

23   programmed means for repeating the removing step until all
24 genes are listed in only one cluster. **Page 8 Line 27-Page 9 Line 7**
25 **(Paragraph 39) last sentence**

1   14. A program product having computer readable code stored on a

2   recordable media for determining similarity between portions of gene

3   expression profiles comprising the steps of:

4   programmed means for processing a number of gene expression

5   profiles with a similar sequences algorithm that is a time and

6   intensity invariant correlation function to obtain a data set of

7   gene expression pairs and a match fraction for each pair; **Page**

8   **7,Line 31-Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**

9   programmed means for choosing a threshold match fraction; **Page**

10   **13 Lines 6-20 (Paragraph 52) sentence 6 and Fig.2 # 217**

11   programmed means for listing gene expression pairs in clusters

12   by their match fractions above the threshold; **Page 8 Lines 16-26**

13   **(paragraph 38) and Fig.2 # 223**

14   programmed means for adding each gene not already in a cluster

15   to a cluster having another gene having a highest match fraction

16   disregarding the threshold with the each gene; **Page 9 Lines 8-14**

17   **(Paragraph 40) and Fig 2 # 229**

18   programmed means for removing a first gene from a first cluster

19   when the first gene is also in a second cluster which has another

20   gene with a higher match fraction with the first gene than any of

21   the genes in the first cluster have with the first gene, **Page 8 Line**

22   **27-Page 9 Line 7 (Paragraph 39) first sentence and Fig. 2 # 225**

23   but;

24   programmed means for retaining the first gene in the first

25   cluster and removing the first gene from the second cluster when the

26   difference between the highest match fraction of the first gene with

27   a gene in the first cluster and the highest match fraction of the

28   first gene with a gene in the second cluster is less than a minimum

29   difference threshold and the number of subsequences represented in

30   the similar gene pair having the highest match fraction in the first

31   cluster is higher than the number of subsequences represented in the

1   similar gene pair having the highest match fraction in the second

2   cluster; **Page 8 Line 27-Page 9 Line 7 (Paragraph 39) second sentence**

3   **and Fig. 2 # 225**

4   programmed means for repeating the removing and retaining steps

5   until all genes are listed in only one cluster. **Page 8 Line 27-Page**

6   **9 Line 7 (Paragraph 39) last sentence**

15. A program product having computer readable code stored on a recordable media for determining similarity between genes comprising the steps of:

programmed means for listing genes to be compared by their gene expression profiles; **Page 7,Line 31-Page 8,Line 9 (Paragraph 36) second sentence and Fig. 1 # 211**

programmed means for processing the listed gene expression profiles with a similar sequences algorithm that is a time and intensity invariant correlation function to obtain a data set of gene expression pairs and a match fraction for each pair; **Page 7,Line 31-Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**

programmed means for choosing a threshold match fraction; **Page 13 Lines 6-20 (Paragraph 52) sentence 6 and Fig.2 # 217**

programmed means for creating a null set G(0) to hold genes accounted for; **Page 13 Lines 6-20 (Paragraph 52) 3rd and 4th sentences and Fig. 2 # 217**

programmed means for assigning genes i and j to cluster 1 if they have a match fraction greater than the threshold; **Page 13 Line 21-Page 14 Line 6 and Page 14 Lines 8-12 (Paragraphs 53 and 54) first sentence and Table IV**

programmed means for assigning gene k to cluster 1 if it has a match fraction greater than the threshold with either gene i or gene j; **Page 13 Line 21-Page 14 Line 6, Page 14 Lines 8-12, Page 14 Line 24-Page 15 Line 3 (Paragraphs 53, 54 2nd sentence, 56) and Table IV**

programmed means for assigning genes k and l to cluster 2 if they have a match fraction greater than the threshold and if both gene k and gene l do not have match fractions above the threshold with either gene i or gene j; **Page 13 Line 21-Page 14 Line 6, Page 14 Lines 8-12, Page 14 Line 24-Page 15 Line 3 (Paragraphs 53, 54 2nd sentence, 56) and Table IV and Fig. 2 # 225**

1   programmed means for removing a first gene from a cluster when
2   another cluster has another gene with a higher match fraction with
3   the first gene; **Page 14 Line 24 – Page 15 Line 3 (Paragraph 56)**
4   **beginning at the 3rd sentence and Fig. 2 # 225**
5   programmed means for repeating the removing step until all
6   genes are listed in only one cluster. **Page 8 Line 27-Page 9 Line 7**
7   **(Paragraph 39) last sentence**
8

16. A program product having computer readable code stored on a recordable media for determining similarity between genes comprising the steps of:

programmed means for listing genes to be compared by their gene expression profiles; **Page 7,Line 31-Page 8,Line 9 (Paragraph 36) second sentence and Fig. 1 # 211**

programmed means for processing the listed gene expression profiles with a similar sequences algorithm that is a time and intensity invariant correlation function to obtain a data set of gene expression pairs and a match fraction for each pair; **Page 7,Line 31-Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**

programmed means for choosing a threshold match fraction; **Page 13 Lines 5-20 (Paragraph 52) sentence 6 and Fig.2 # 217**

programmed means for creating a null set G(0) to hold genes accounted for; **Page 8 Lines 11-14 (Paragraph 37 and Fig. 2 # 217**

programmed means for assigning genes i and j to cluster 1 if they have a match fraction greater than the threshold; **Page 13 Line 21-Page 14 Line 6 and Page 14 Lines 8-12 (Paragraphs 53 and 54) first sentence and Table IV**

programmed means for assigning gene k to cluster 1 if it has a match fraction greater than the threshold with either gene i or gene j; **Page 13 Line 21-Page 14 Line 6, Page 14 Lines 8-12, Page 14 Line 24-Page 15 Line 3 (Paragraphs 53, 54 2nd sentence, 56) and Table IV**

programmed means for assigning genes k and l to cluster 2 if they have a match fraction greater than the threshold and if both gene k and gene l do not have match fractions above the threshold with either gene i or gene j; **Page 13 Line 21-Page 14 Line 6, Page 14 Lines 8-12, Page 14 Line 24-Page 15 Line 3 (Paragraphs 53, 54 2nd sentence, 56) and Table IV and Fig. 2 # 225**

1   programmed means for removing a first gene from a cluster when
2 another cluster has another gene with a higher match fraction with
3 the first gene, unless the another gene requires a larger number of
4 subsequences to achieve similarity with the first gene; **Page 14 Line**
5 **24 - Page 15 Line 3 (Paragraph 56) beginning at the 3rd sentence and**
6 **Fig. 2 # 225**
7   programmed means for repeating the removing step until all
8 genes are listed in only one cluster. **Page 8 Line 27-Page 9 Line 7**
9 **(Paragraph 39) last sentence**

1   20. In a method of determining similarity between portions of
2   gene expression profiles which includes processing a number of gene
3   expression profiles using a computer with a similar sequences
4   algorithm that is a time and intensity invariant correlation
5   function to obtain a data set of gene expression pairs and a match
6   fraction for each pair, **Page 7,Line 31-Page 8,Line 9 (Paragraph 36**
7   **and Fig. 1, # 213, 215** the improvement comprising the steps of:
8   listing gene expression pairs in clusters by their match
9   fractions; **Page 8 Lines 16-26 (paragraph 38) and Fig.2 # 223**
10   removing a first gene from a cluster when another cluster has
11   another gene with a higher match fraction with the first gene,
12   unless the another gene requires a larger number of subsequences to
13   achieve similarity with the first gene; **Page 14 Line 24 – Page 15**
14   **Line 3 (Paragraph 56) beginning at the 3rd sentence and Fig. 2 # 225**
15   repeating the removing step until all genes are listed in only
16   one cluster; **Page 8 Line 27-Page 9 Line 7 (Paragraph 39) last**
17   **sentence**
18   providing output of the listing of clusters of gene
19   expression profiles. **Page 6 Lines 24-31 and Page 7 Lines 19-24**
20   **(Paragraphs 32 and 34) and Fig. 8 #s 54 and 66**
21
22
23
24

(VI) Grounds of Rejection to be reviewed on Appeal

Claims 1-6, 10-16 and 20 are pending in this application.

Claims 7-9 and 17-19 were canceled after restriction.

The rejection of claims 10-16 and 20 under 35 U.S.C. 101 and
the rejection of claims 1-6 under 35 U.S.C. 101 and for new matter
is appealed.

The issues in this appeal are:

I.   Whether output to a user is a required claim step in order to
define an invention, that is a practical application which is
useful, concrete and tangible.

Claim 10 is representative of claims 10 - 16 which have been
rejected under 35 U.S.C. 101 and is related to Issue I

II.   Whether applicants' teaching of a personal computer with
implicit, intrinsic and inherent output means in the specification
support claims 1 - 6 and 20 without adding new matter.

Claim 1 is representative of claims 1 - 6, 20 which have been
rejected as containing new matter and is related to Issue II

1   <u>**(VII) Argument**</u>

3   **Issue I:** Whether output to a user is a required claim step in order
4   to define an invention, that is a practical application which is
5   useful, concrete and tangible.

7   Appellants claim in representative claim 10:
8   10. A program product having computer readable code stored on a
9   recordable media for determining similarity between portions of gene
10   expression profiles comprising:
11   programmed means for processing a number of gene expression
12   profiles with a similar sequences algorithm that is a time and
13   intensity invariant correlation function to obtain a data set of
14   gene expression pairs and a match fraction for each pair; **Page
15   7,Line 31-Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**
16   programmed means for listing gene expression pairs in clusters
17   by their match fractions; **Page 8 Lines 16-26 (paragraph 38) and
18   Fig.2 # 223**
19   programmed means for removing a first gene from a cluster when
20   the first gene is also in another cluster which has another gene
21   with a higher match fraction with the first gene than any of the
22   genes in the cluster have with the first gene; **Page 8 Line 27-Page 9
23   Line 7 (paragraph 39) and Fig. 2 # 225**
24   programmed means for repeating the removing step until all
25   genes are listed in only one cluster. **Page 8 Line 27-Page 9 Line 7
26   (Paragraph 39) last sentence**

Applicants believe that the specification and claims indeed do describe a method and a program product that produce a result that has substantial and credible utility as required by MPEP 2107 II and that the claims are limited to a narrow practical application in a computer related art.

The Examiner relies on the "New Interim Guidelines" to interpret the requirements of the Federal Courts under the current law to require claiming "output to a user". Applicants believe that the Examiner is mistaken and is applying an interpretation of the definition of the word tangible that is:
1) narrower than appropriate under the current law and is
2) narrower than required under the "New Guidelines".

1) The introduction to the "New Guidelines" states:
"These Examination Guidelines ("Guidelines") are based on the USPTO's current understanding of the law and are believed to be fully consistent with binding precedent of the Supreme Court, the Federal Circuit and the Federal Circuit's predecessor courts. These Guidelines do not constitute substantive rulemaking and hence do not have the force and effect of law."

In following the "Guidelines", the Examiner appears to require separate interpretations of the words useful, concrete and tangible.

Applicants' attorney has found no basis in any of the Federal Circuit opinions using these words that imply that these terms are to have separate meanings. They appear to always be used together as synonyms for the concept of being useful and non-abstract. Applicants' attorney has requested that the Examiner provide a citation to a court's requirement that these terms are part of a three pronged test if such is the case in order to help applicants

decide whether to appeal or request continued examination. No citation was provided.

2) Even under the "Guidelines, the Examiners interpretation of the word tangible is unnecessarily narrow.
The "Guidelines" at page 13 recite

"Accordingly, a complete definition of the scope of 35 U.S.C. § 101, reflecting Congressional intent, is that any new and useful process, machine, manufacture or composition of matter under the sun that is made by man is the proper subject matter of a patent. The subject matter courts have found to be outside of, or exceptions to, the four statutory categories of invention is limited to abstract ideas, laws of nature and natural phenomena. While this is easily stated, determining whether an applicant is seeking to patent an abstract idea, a law of nature or a natural phenomenon has proven to be challenging."

Beginning at page 21 the "Guidelines" recite:

"TANGIBLE RESULT"

"The tangible requirement does not necessarily mean that a claim must either be tied to a particular machine or apparatus or must operate to change articles or materials to a different state or thing. However, the tangible requirement does require that the claim must recite more than a § 101 judicial exception, in that the process claim must set forth a practical application of that § 101 judicial exception to produce a real-world result. Benson, 409 U.S. at 71-72, 175 USPQ at 676-77 (invention ineligible because had "no substantial practical application."). "[A]n application of a law of nature or mathematical formula to a ... process may well be deserving of patent protection." Diehr, 450 U.S. at 187, 209 USPQ at 8 (emphasis added); see also Corning, 56 U.S. (15 How.) at 268, 14 L.Ed. 683 ("It is for the discovery or invention of some practical method or means of producing a beneficial result or effect, that a patent is granted ...").

In other words, the opposite meaning of "tangible" is "abstract."
The bare conversion of any binary data as in Gottschalk V. Benson or the bubble sort of any data as in "Warmerdam, 33 F.3d at 1360, 31 USPQ2d at 1759 ("steps of 'locating' a medial axis, and 'creating' a bubble hierarchy ... describe nothing more than the

manipulation of basic mathematical constructs, the paradigmatic 'abstract idea'")" recited at page 14 of the "Guidelines" are examples of the abstract.

Applicants' process does not convert or process just any data but is limited to useful concrete and non-abstract gene expression profiles in a data base of such profiles. Applicants' process is but one application of many possible applications of the mathematical steps involved in obtaining the useful result.

At page 17 of the "Guidelines we see:

While abstract ideas, natural phenomena, and laws of nature are not eligible for patenting, methods and products employing abstract ideas, natural phenomena, and laws of nature to perform a real-world function may well be. In evaluating whether a claim meets the requirements of section 101, the claim must be considered as a whole to determine whether it is for a particular application of an abstract idea, natural phenomenon, or law of nature, rather than for the abstract idea, natural phenomenon, or law of nature itself.

As is clear from the specification and the claim limitations, applicants' process is limited to a particular practical application and is not an abstract idea, natural phenomenon or a law of nature.

The result is that all of the processed gene expression profiles are each listed in only one cluster. This result of applicants' claims is a very useful, repeatable and non-abstract result which is recognized by those skilled in the medical and computer arts to be of great value and useful, non-abstract and concrete finding of similar gene expression profiles.

PRIOR ART

Applicants note that their claims have not been rejected on prior art yet have been restricted on the ground that there were two groups of claims that required two fields of search. It is not apparent whether relevant prior art patents were considered by the Examiner while examining this application. It is believed that the "Guidelines" on page 10 are helpful in determining both the novelty of applicants' invention and the **usefulness and non-abstract nature** of applicants' the invention.

As evidenced by the references which applicants have attempted to incorporate by reference, but have acquiesced to the Examiners correct requirement to cancel, in addition to applicants teachings in the background art section of their specification, users in the medical profession find great value and usefulness in methods for finding similar gene expression profiles that are tangible and concrete. See for example US Patent 6,406,853 abstract and claims 25, 26 and US Patent 6,436,642 column 26 beginning at line 15.

It is believed that if the rejections under 35 U.S.C. 101 put forth in this application were appropriate, many of the relevant prior art patents in the appropriate fields of search would be found to be invalid. Since they were issued under the guidance of current statutory law and court cases, it must be that the rejections in this application are based upon excessively narrow and untenable interpretation of the current law.

1   **Issue II**: Whether applicants' teaching of a personal computer
2   with implicit, intrinsic and inherent output means in the
3   specification support claims 1 - 6 without adding new matter.
4
5   Appellants claim in representative claim 1:
6   1. A method for determining similarity between portions of gene
7   expression profiles in a computer comprising the steps of:
8   processing a number of gene expression profiles with a similar
9   sequences algorithm that is a time and intensity invariant
10   correlation function to obtain a data set of gene expression profile
11   pairs and a match fraction for each gene expression profile pair;
12   **Page 7,Line 31-Page 8,Line 9 (Paragraph 36) and Fig. 1, # 213, 215**
13   listing gene expression profile pairs in clusters by their
14   match fractions; **Page 8 Lines 16-26 (paragraph 38) and Fig.2 # 223**
15   removing a first gene expression profile from a cluster when
16   another cluster has another gene expression profile with a higher
17   match fraction with the first gene expression profile, unless the
18   another gene expression profile requires a larger number of
19   subsequences to achieve similarity with the first gene expression
20   profile; **Page 8 Line 27-Page 9 Line 7 (paragraph 39) and Fig. 2 #**
21   **225**
22   repeating the removing step until all gene expression profiles
23   are listed in only one cluster; **Page 8 Line 27-Page 9 Line 7**
24   **(Paragraph 39) last sentence**
25   providing output of the listing of clusters of gene
26   expression profiles. **Page 6 Lines 24-31 and Page 7 Lines 19-24**
27   **(Paragraphs 32 and 34) and Fig. 8 #s 54 and 66**
28

Applicants' specification recites: The focal point of the preferred personal computer architecture comprises a processor 51. The processor 51 is connected to a bus 52 which comprises a set of data lines, a set of address lines and a set of control lines. A plurality of I/O devices, memory and storage devices 53-58 and 66 are connected to the bus 52 through separate adapters 59-64 and 67, respectively. For example, the display 54 may be either a CRT or a flat panel display.

It is believed to be well known in the art as exemplified by prior art patents that users in the medical profession receive output from personal computer input/output devices such as applicants teach in their preferred embodiment. Again, applicants refer to US Patent 6,406,853 abstract and claims 25, 26 and US Patent 6,436,642 column 26 beginning at line 15.

It is believed that material that is implicit, intrinsic, or inherent in the application as filed is not new matter.

In order to be usable by a user, a personal computer necessarily and constantly exhibits the function of input and output, and such function was recognized as such by those skilled in the art of using personal computers. Therefore applicants' addition of the step of providing such output to satisfy the Examiner's reading of the guidelines was not new matter but is supported in their specification by teachings that are implicit, intrinsic and inherent.

Accordingly it is believed that the claims are clear, statutory and definite and are drawn to a novel and unobvious method and program product for clustering gene expression profiles which result is concrete, tangible and directly useful in drug selection and disease diagnosis.

Request for Relief
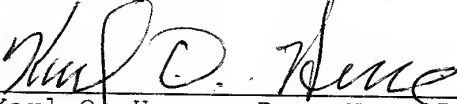
Wherefore, Appellants respectfully request that the rejection of pending claims 1 - 6, 10 - 16 and 20 be reversed.

Respectfully submitted,

Date: October 3, 2007

Karl O. Hesse, Reg. No. 25,398
Attorney for Appellants

IBM Corporation
Intellectual Property Law
MG90-201/1
8501 IBM Drive
Charlotte, NC 28262-8563

Land line (704) 895-8241
Cell phone (704) 724-1413
Fax: (704) 594-8307

(VIII) **Appendix**

Claims Involved in this Appeal

1. A method for determining similarity between portions of gene expression profiles in a computer comprising the steps of:

processing a number of gene expression profiles with a similar sequences algorithm that is a time and intensity invariant correlation function to obtain a data set of gene expression profile pairs and a match fraction for each gene expression profile pair;

listing gene expression profile pairs in clusters by their match fractions;

removing a first gene expression profile from a cluster when another cluster has another gene expression profile with a higher match fraction with the first gene expression profile, unless the another gene expression profile requires a larger number of subsequences to achieve similarity with the first gene expression profile;

repeating the removing step until all gene expression profiles are listed in only one cluster;

providing output of the listing of clusters of gene expression profiles.

1   2. A method for determining similarity between portions of gene
2 expression profiles comprising the steps of:
3   processing a number of gene expression profiles with a similar
4 sequences algorithm that is a time and intensity invariant
5 correlation function to obtain a data set of gene expression pairs
6 and a match fraction for each pair;
7   listing gene expression pairs in clusters by their match
8 fractions;
9   removing a first gene from a first cluster when the first gene
10 is also in a second cluster which has another gene with a higher
11 match fraction with the first gene than any of the genes in the
12 first cluster have with the first gene, but;
13   retaining the first gene in the first cluster and removing the
14 first gene from the second cluster when the difference between the
15 highest match fraction of the first gene with a gene in the first
16 cluster and the highest match fraction of the first gene with a gene
17 in the second cluster is less than a minimum difference threshold
18 and the number of subsequences represented in the similar gene pair
19 having the highest match fraction in the first cluster is higher
20 than the number of subsequences represented in the similar gene pair
21 having the highest match fraction in the second cluster;
22   repeating the removing step until all genes are listed in only
23 one cluster;
24   providing output of the listing of clusters of gene
25 expression profiles.

1   3. A method of determining similarity between portions of gene
2   expression profiles comprising the steps of:
3   processing data embodying a number of gene expression profiles
4   with a similar sequences algorithm in a computer that is a time and
5   intensity invariant correlation function to obtain a data set of
6   gene expression pairs and a match fraction for each pair;
7   choosing a threshold match fraction;
8   listing gene expression pairs in clusters by their match
9   fractions above the threshold;
10   adding each gene not already in a cluster to a cluster having
11   another gene having a highest match fraction with the each gene
12   without regard of the threshold;
13   removing a first gene from a cluster when the first gene is
14   also in another cluster which has another gene with a higher match
15   fraction with the first gene than any of the genes in the cluster
16   have with the first gene;
17   repeating the removing step until all genes are listed in only
18   one cluster;
19   providing output of the listing of clusters of gene
20   expression profiles.

1   4. A method for determining similarity between portions of gene
2 expression profiles comprising the steps of:
3   processing a number of gene expression profiles with a similar
4 sequences algorithm that is a time and intensity invariant
5 correlation function with a computer to obtain a data set of gene
6 expression pairs and a match fraction for each pair;
7   choosing a threshold match fraction;
8   listing gene expression pairs in clusters by their match
9 fractions above the threshold;
10   adding each gene not already in a cluster to a cluster having
11 another gene having a highest match fraction disregarding the
12 threshold with the each gene;
13   removing a first gene from a first cluster when the first gene
14 is also in a second cluster which has another gene with a higher
15 match fraction with the first gene than any of the genes in the
16 first cluster have with the first gene, but;
17   retaining the first gene in the first cluster and removing the
18 first gene from the second cluster when the difference between the
19 highest match fraction of the first gene with a gene in the first
20 cluster and the highest match fraction of the first gene with a gene
21 in the second cluster is less than a minimum difference threshold
22 and the number of subsequences represented in the similar gene pair
23 having the highest match fraction in the first cluster is higher
24 than the number of subsequences represented in the similar gene pair
25 having the highest match fraction in the second cluster;
26   repeating the removing and retaining steps until all genes are
27 listed in only one cluster;
28   providing output of the listing of clusters of gene
29 expression profiles.

5. A method in a computer for determining similarity between genes comprising the steps of:

listing genes to be compared in a data set by their gene expression profiles;

processing the listed gene expression profiles with a similar sequences algorithm that is a time and intensity invariant correlation function to obtain a data set of gene expression pairs and a match fraction for each pair;

choosing a threshold match fraction;

creating a set G in which to list indices of genes accounted for;

assigning genes i and j to a cluster a if they have a match fraction greater than the threshold;

assigning gene k to the cluster a if it has a match fraction greater than the threshold with either gene i or gene j;

assigning genes k and l to a cluster b if they have a match fraction greater than the threshold and if both gene k and gene l do not have match fractions above the threshold with either gene i or gene j;

repeating the assigning steps until all genes to be compared have been considered;

removing a first gene from a cluster when another cluster has another gene with a higher match fraction with the first gene;

repeating the removing step until all genes are listed in only one cluster;

providing output of the listing of clusters of gene expression profiles.

6. A method in a computer for determining similarity between genes comprising the steps of:

listing genes to be compared in a data set by their gene expression profiles;

processing the listed gene expression profiles with a similar sequences algorithm that is a time and intensity invariant correlation function to obtain a data set of gene expression pairs and a match fraction for each pair;

choosing a threshold match fraction;

creating a set G in which to list indices of genes accounted for;

assigning genes i and j to cluster 1 if they have a match fraction greater than the threshold;

assigning gene k to cluster 1 if it has a match fraction greater than the threshold with either gene i or gene j;

assigning genes k and l to cluster 2 if they have a match fraction greater than the threshold and if both gene k and gene l do not have match fractions above the threshold with either gene i or gene j;

removing a first gene from a cluster when another cluster has another gene with a higher match fraction with the first gene, unless the another gene requires a larger number of subsequences to achieve similarity with the first gene;

repeating the removing step until all genes are listed in only one cluster;

providing output of the listing of clusters of gene expression profiles.

1   10. A program product having computer readable code stored on a
2   recordable media for determining similarity between portions of gene
3   expression profiles comprising:
4   programmed means for processing a number of gene expression
5   profiles with a similar sequences algorithm that is a time and
6   intensity invariant correlation function to obtain a data set of
7   gene expression pairs and a match fraction for each pair;
8   programmed means for listing gene expression pairs in clusters
9   by their match fractions;
10   programmed means for removing a first gene from a cluster when
11   the first gene is also in another cluster which has another gene
12   with a higher match fraction with the first gene than any of the
13   genes in the cluster have with the first gene;
14   programmed means for repeating the removing step until all
15   genes are listed in only one cluster.

1   11. A program product having computer readable code stored on a
2   recordable media for determining similarity between portions of gene
3   expression profiles using output from a similar sequences algorithm
4   that is a time and intensity invariant correlation function
5   comprising:
6   programmed means for providing a gene expression profile data
7   set as input to programmed means embodying a similar sequences
8   algorithm that is a time and intensity invariant correlation
9   function to obtain a data set of gene expression pairs and a match
10   fraction for each pair as output from the programmed means embodying
11   a similar sequences algorithm;
12   programmed means for listing the gene expression pairs in
13   clusters by their match fractions;
14   programmed means for removing a first gene from a cluster when
15   the first gene is also in another cluster which has another gene
16   with a higher match fraction with the first gene than any of the
17   genes in the cluster have with the first gene;
18   programmed means for repeating the removing step until all
19   genes are listed in only one cluster.

1   12. A program product having computer readable code stored on a

2   recordable media for determining similarity between portions of gene

3   expression profiles comprising the steps of:

4   programmed means for processing a number of gene expression

5   profiles with a similar sequences algorithm that is a time and

6   intensity invariant correlation function to obtain a data set of

7   gene expression pairs and a match fraction for each pair;

8   programmed means for listing gene expression pairs in clusters

9   by their match fractions;

10   programmed means for removing a first gene from a first cluster

11   when the first gene is also in a second cluster which has another

12   gene with a higher match fraction with the first gene than any of

13   the genes in the first cluster have with the first gene, but;

14   programmed means for retaining the first gene in the first

15   cluster and removing the first gene from the second cluster when the

16   difference between the highest match fraction of the first gene with

17   a gene in the first cluster and the highest match fraction of the

18   first gene with a gene in the second cluster is less than a minimum

19   difference threshold and the number of subsequences represented in

20   the similar gene pair having the highest match fraction in the first

21   cluster is higher than the number of subsequences represented in the

22   similar gene pair having the highest match fraction in the second

23   cluster;

24   programmed means for repeating the removing step until all

25   genes are listed in only one cluster.

1   13. A program product having computer readable code stored on a
2 recordable media for determining similarity between portions of gene
3 expression profiles comprising the steps of:
4   programmed means for processing a number of gene expression
5 profiles with a similar sequences algorithm that is a time and
6 intensity invariant correlation function to obtain a data set of
7 gene expression pairs and a match fraction for each pair;
8   programmed means for choosing a threshold match fraction;
9   programmed means for listing gene expression pairs in clusters
10 by their match fractions above the threshold;
11   programmed means for adding each gene not already in a cluster
12 to a cluster having another gene having a highest match fraction
13 with the each gene without regard of the threshold;
14   programmed means for removing a first gene from a cluster when
15 the first gene is also in another cluster which has another gene
16 with a higher match fraction with the first gene than any of the
17 genes in the cluster have with the first gene;
18   programmed means for repeating the removing step until all
19 genes are listed in only one cluster.

1   14. A program product having computer readable code stored on a
2   recordable media for determining similarity between portions of gene
3   expression profiles comprising the steps of:
4   programmed means for processing a number of gene expression
5   profiles with a similar sequences algorithm that is a time and
6   intensity invariant correlation function to obtain a data set of
7   gene expression pairs and a match fraction for each pair;
8   programmed means for choosing a threshold match fraction;
9   programmed means for listing gene expression pairs in clusters
10   by their match fractions above the threshold;
11   programmed means for adding each gene not already in a cluster
12   to a cluster having another gene having a highest match fraction
13   disregarding the threshold with the each gene;
14   programmed means for removing a first gene from a first cluster
15   when the first gene is also in a second cluster which has another
16   gene with a higher match fraction with the first gene than any of
17   the genes in the first cluster have with the first gene, but;
18   programmed means for retaining the first gene in the first
19   cluster and removing the first gene from the second cluster when the
20   difference between the highest match fraction of the first gene with
21   a gene in the first cluster and the highest match fraction of the
22   first gene with a gene in the second cluster is less than a minimum
23   difference threshold and the number of subsequences represented in
24   the similar gene pair having the highest match fraction in the first
25   cluster is higher than the number of subsequences represented in the
26   similar gene pair having the highest match fraction in the second
27   cluster;
28   programmed means for repeating the removing and retaining steps
29   until all genes are listed in only one cluster.

15. A program product having computer readable code stored on a recordable media for determining similarity between genes comprising the steps of:

programmed means for listing genes to be compared by their gene expression profiles;

programmed means for processing the listed gene expression profiles with a similar sequences algorithm that is a time and intensity invariant correlation function to obtain a data set of gene expression pairs and a match fraction for each pair;

programmed means for choosing a threshold match fraction;

programmed means for creating a null set $G(0)$ to hold genes accounted for;

programmed means for assigning genes i and j to cluster 1 if they have a match fraction greater than the threshold;

programmed means for assigning gene k to cluster 1 if it has a match fraction greater than the threshold with either gene i or gene j;

programmed means for assigning genes k and l to cluster 2 if they have a match fraction greater than the threshold and if both gene k and gene l do not have match fractions above the threshold with either gene i or gene j;

programmed means for removing a first gene from a cluster when another cluster has another gene with a higher match fraction with the first gene;

programmed means for repeating the removing step until all genes are listed in only one cluster.

16. A program product having computer readable code stored on a recordable media for determining similarity between genes comprising the steps of:

programmed means for listing genes to be compared by their gene expression profiles;

programmed means for processing the listed gene expression profiles with a similar sequences algorithm that is a time and intensity invariant correlation function to obtain a data set of gene expression pairs and a match fraction for each pair;

programmed means for choosing a threshold match fraction;

programmed means for creating a null set $G(0)$ to hold genes accounted for;

programmed means for assigning genes i and j to cluster 1 if they have a match fraction greater than the threshold;

programmed means for assigning gene k to cluster 1 if it has a match fraction greater than the threshold with either gene i or gene j;

programmed means for assigning genes k and l to cluster 2 if they have a match fraction greater than the threshold and if both gene k and gene l do not have match fractions above the threshold with either gene i or gene j;

programmed means for removing a first gene from a cluster when another cluster has another gene with a higher match fraction with the first gene, unless the another gene requires a larger number of subsequences to achieve similarity with the first gene;

programmed means for repeating the removing step until all genes are listed in only one cluster.

1   20. In a method of determining similarity between portions of
2   gene expression profiles which includes processing a number of gene
3   expression profiles using a computer with a similar sequences
4   algorithm that is a time and intensity invariant correlation
5   function to obtain a data set of gene expression pairs and a match
6   fraction for each pair, the improvement comprising the steps of:
7   listing gene expression pairs in clusters by their match
8   fractions;
9   removing a first gene from a cluster when another cluster has
10   another gene with a higher match fraction with the first gene,
11   unless the another gene requires a larger number of subsequences to
12   achieve similarity with the first gene;
13   repeating the removing step until all genes are listed in only
14   one cluster;
15   providing output of the listing of clusters of gene
16   expression profiles.
17

1   **(IX) Evidence Appendix**

2   No evidence is being submitted in this appeal.

3

1   **(X) Related Proceedings Appendix**

2   None.